DOCUMENT RESUME

ED 107 712 95 TM 004 553

AUTHOR Starr, B. James
TITLE Psychosocial Maturity: A Preliminary Examination of Validation Techniques. Report No. 157.
INSTITUTION Johns Hopkins Univ., Baltimore, Md. Center for the Study of Social Organization of Schools.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO R-157
PUB DATE Jun 73
CONTRACT NE-C-00-3-0113
NOTE 45p.; Appendices C and D are illegible but do not detract significantly from the document

EDRS PRICE MF-$0.76 HC-$1.95 PLUS POSTAGE
DESCRIPTORS Elementary Education; Interpersonal Competence; Matrices; *Maturity Tests; Measurement Techniques; Peer Relationship; Predictive Ability (Testing); *Predictive Validity; *Rating Scales; *Social Maturity; Statistical Analysis; Test Reliability; *Test Validity
IDENTIFIERS *Psychosocial Maturity Scale

ABSTRACT

Employed were a small sample of fifth graders to examine techniques for assessing the validity of psychosocial maturity (PSM) and the scales used to measure it. Three validation techniques were examined: (1) Data obtained from self-report scales, teacher ratings, and peer ratings were used to generate three multitrait-multimethod matrices (total sample, boys, and girls), and the criteria outlined by Campbell and Fiske (1959) were applied to examine these matrices; (2) a technique was devised to assess the degree to which independent raters could predict the pattern of relationships among PSM and other scale scores; and (3) PSM subscales were used to generate regression coefficients predicting rated traits. Validity studies in general and some implications for the validation of a new PSM scale are discussed. A major implication for future studies is that validation of component subscales of PSM, rather than of the multidimensional construct as a whole, should be attempted. (Author)

Center for Social Organization of Schools

Report No. 157
June, 1973

BEST COPY AVAILABLE

PSYCHOSOCIAL MATURITY: A PRELIMINARY EXAMINATION OF VALIDATION TECHNIQUES

B. James Starr

The Johns Hopkins University

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

STAFF

John L. Holland, Director

James M. McPartland, Assistant Director

Joan E. Brown

Judith P. Clark

David L. DeVries

Keith J. Edwards

Gail M. Fennessey

James J. Fennessey

Stephanie G. Freeman

Ellen Greenberger

Edward J. Harsch

Samuel T. Helms

John H. Hollifield

Patricia A. Hughes

Nancy L. Karweit

Samuel A. Livingston

Edward McDill

Alyce J. Nafziger

Dean H. Nafziger

Karen A. Schwartzman

John P. Snyder

Julian C. Stanley

Gerald D. Williams

ED107712

PSYCHOSOCIAL MATURITY: A PRELIMINARY EXAMINATION OF VALIDATION TECHNIQUES

CONTRACT NO. NE-C-00-3-0113

WORK UNIT 2

B. JAMES STARR

REPORT NO. 157

June, 1973

Published by the Center for Social Organization of Schools, supported in part as a research and development center by funds from the United States National Institute of Education, Department of Health, Education, and Welfare. The opinions expressed in this publication do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by the Institute should be inferred.

The Johns Hopkins University

Baltimore, Maryland

INTRODUCTORY STATEMENT

The Center for Social Organization of Schools has two primary objectives: to develop a scientific knowledge of how schools affect their students, and to use this knowledge to develop better school practices and organization.

The Center works through three programs to achieve its objectives. The Schools and Maturity program is studying the effects of school, family, and peer group experiences on the development of attitudes consistent with psychosocial maturity. The objectives are to formulate, assess, and research important educational goals other than traditional academic achievement. The School Organization program is currently concerned with authority-control structures, task structures, reward systems, and peer group processes in schools. The Careers and Curricula program bases its work upon a theory of career development. It has developed a self-administered vocational guidance device and a self-directed career program to promote vocational development and to foster satisfying curricular decisions for high school, college, and adult populations.

This report, prepared by the Schools and Maturity program, examines three techniques that may be useful in validating the psychosocial maturity scale that is currently being revised by the program.

ABSTRACT

The present investigation employed a small sample of fifth graders to examine techniques for assessing the validity of psychosocial maturity (PSM) and the scales used to measure it. Three validation techniques were examined: (1) Data obtained from self-report scales, teacher ratings, and peer ratings were used to generate three multitrait-multimethod matrices (total sample, boys, and girls), and the criteria outlined by Campbell and Fiske (1959) were applied to examine these matrices; (2) a technique was devised to assess the degree to which independent raters could predict the pattern of relationships among PSM and other scale scores; and (3) PSM subscales were used to generate regression coefficients predicting rated traits.

Validity studies in general and some implications for the validation of a new PSM scale are discussed. A major implication for future studies is that validation of component subscales of PSM, rather than of the multidimensional construct as a whole, should be attempted.

ACKNOWLEDGMENTS

The author would like to express his gratitude to the administration of the Friends School in Baltimore for their cooperation in the conducting of this study. I especially thank the teachers, Mrs. Jeanine Bitner and Mrs. Claire Loecher, for the use of their class time and their contribution as raters in this study. A special debt is owed to Ellen Greenberger and Joan Bitner Brown who made contributions in all phases of this investigation. I would also like to express my thanks to Alex Seidler for help in developing the pattern-matching technique employed here and for his helpful responses to earlier drafts of this report. Finally, I would like to thank Samuel Livingston, who also read and offered helpful suggestions on drafts of the paper.

INTRODUCTION

A recent paper by Greenberger and Sørensen (1971) provided the theoretical groundwork for evolving a measure of psychosocial maturity (PSM). The maturity concept developed in that paper focused on the skills and attitudes thought necessary to promote effective functioning in society. Employing an interdisciplinary model, the concept encompassed biological, sociological, and psychological aspects of maturity and outlined three capacities of the mature individual: effective individual functioning, effective interpersonal functioning, and capacity to contribute to maintenance of the social system.

The multifactor scale designed to assess PSM in American society was reported by Greenberger et al (1971). Items were taken from an item-pool administered independently and for other purposes by the Pennsylvania State Department of Education. Items were selected on the basis of their theoretical relevance and their differential endorsement by a large sample of 5th and 11th graders.[1] The items so chosen were subjected to a 5-factor Principal Components solution, the number of factors having been selected according to previous theoretical groupings. The resulting factors (subscales) were called self-esteem, openness to change, independence, identity, and social tolerance.

Since the development of the scale, a number of substantive studies have focused on PSM (e.g., Greenberger and Marini, 1972; Starr et al, 1972). Evidence for the validity of the scale has derived, in a largely incidental way, from its behavior in these investigations. One recent study (Greenberger, 1972) has come to grips with the validity issue a bit more directly. This study examined the question of whether or not the PSM

[1] The latter criterion reflects the minimal logical consistency with a notion of temporal growth.

scale simply measured social desirability. Results of that study demonstrated that two related measures of social desirability yielded only minimal correlations with PSM scores. No other study has, to date, dealt directly with the validity problem.

The PSM scale is currently being revised. Nonetheless, a study examining the validity of the old scale should be useful for exploring various procedures to validate the newly-designed scale. The present study employs a variety of techniques for assessing the predictive validity of the scale and also attempts to examine construct validity. Here, again, the exploratory function of this research is important in assessing which of two alternative methods of measurement might be more effective (when paired with measurement based on the PSM scale) in defining the construct via a multitrait-multimethod analysis (Campbell and Fiske, 1959).

It is necessary to state, at the outset, some of the major limitations of this investigation. First, although PSM is a measure which shows increments with age, an age-restricted sample was employed here. A readily available sample of fifth graders constituted the present sample. While this truncation of range is admittedly not desirable,[1] it was felt that some of the important questions raised could be answered in terms of the relative magnitudes of the correlation coefficients. Second, none of the raters who produced data for the multitrait-multimethod matrix were trained. It might be anticipated that even a moderate degree of training would elevate the magnitude of measures of relationships generated on the basis of such ratings. Third, the size of the sample was quite small and the coefficients obtained, therefore, are correspondingly unstable. Thus, clearly, the present investigation is not a definitive attempt to assess validity; however,

[1] The truncation of range would be expected to produce spuriously low correlation coefficients.

such an assessment should provide some concrete guidelines for future validity studies.

With these cautions in mind, it is still profitable to outline the idealized expectation. It is hypothesized that the PSM score will correlate well with a measure of social solidarity and progressively more poorly with measures of social desirability and creative tendency. In addition, other (intuitively generated) evidence for the predictive validity of the scale will be examined.

METHOD

Subjects

The subjects were 47 5th graders attending two classes in a private school in Baltimore. This total included 22 girls and 25 boys. It represents all 5th graders present on two consecutive testing days in early June.

Validation techniques

Three validation techniques were examined for their usefulness in assessing the validity of PSM. First, three rating methods (self-report scales, peer ratings, and teacher ratings) and four traits (PSM, social solidarity, social desirability, and creative tendency) were used to form a 12 x 12 matrix of intercorrelations for use in multitrait-multimethod analysis. Second, the degree to which independent raters could predict the pattern of relationships among scale scores was analyzed. Third, PSM subscales were used to generate regression coefficients predicting rated traits.

1) Multitrait-Multimethod Analysis: The traits and ratings used to form the multitrait-multimethod matrix were devised as follows:

Traits. Four scale measures were chosen to represent traits for

the multitrait-multimethod matrix. The PSM[1] construct was the validation target; and hence, other measures will be described as convergent or discriminant with regard to PSM. A social solidarity scale[2], evolved by the Schools and Maturity Program as a measure of feelings of community with others, was included as a convergent trait. Originally, this scale was designed for possible inclusion in the revised PSM scale, to form part of the evaluation of capacity to contribute to system maintenance. Items on this scale are presented in a true-false format. Sample items (with the appropriate "socially solid" answers in parentheses) include: "It's none of my business if my neighbors are in trouble and need help" (-), and "I would not mind serving on a committee for my school." (+)

A modified Crowne-Marlowe (1960) social desirability scale[3] was expected to represent a divergent trait. Finally, a scale of creative tendency[4], developed by the Pennsylvania State Department of Education as part of their Educational Quality Assessment (EQA) program, was included as a more discriminant construct. Responses to items on this scale indicate agreement or disagreement on a 5-point scale. Agreement with items like, "New places are fun to visit," and disagreement with items such as, "Pictures of grass should be colored only green" contribute to an overall "creative" score. There is no theoretical or empirical evidence to suggest that creativity and maturity, as measured by the PSM scale, should be related.

Ratings. Forms for peer ratings and teacher ratings of the selected traits were devised by members of the Schools and Maturity Program. One

---

1 The PSM scale is shown in Appendix A.

2 This scale appears in Appendix B.

3 The modified scale appears in Appendix C.

4 The creative tendency scale is shown in Appendix D.

of the teacher-rating forms is shown in Figure 1; its purpose is to assess creative tendency. The description of an individual with high creative tendency was:

This person:
- has ideas which are often very unusual or surprising
- has talents (for example, writing or telling stories, painting, or carving) that many other children don't have
- sometimes surprises the class with good, but unexpected answers
- has his or her own way of doing things

The description for a high degree of social desirability was:

This person:
- is the kind of person who thinks it is very, very important for other people to like him (her)
- does what he (she) thinks other people would like him (her) to do, rather than what he (she) would like to do
- always tries to do what he or she thinks that other people would say is the "right" thing
- very much likes to please the teacher

The corresponding description for the social solidarity trait was:

This person:
- is the kind of person you would want to have join in on a class project; gives up time to help other kids even if he or she won't get anything out of it
- is not selfish: likes to see other people get what they want or need
- works well with other children
- is friendly to other children, even if he or she doesn't know them very well

Finally, the composite trait description for PSM was:

Who would get along best if suddenly moved far off to some imaginary planet?
No specific talent would be needed there, but the person would need to:
- get along well on his own: pick up useful information, make decisions by himself
- know how to get along with others: act in a way that others can understand; figure out who to trust and how much
- take an interest in the society: learn the customs and values, cooperate with other people

The rating forms were compiled in a booklet with a cover letter broadly explaining the nature of the study to the teachers.[1]

[1] The cover letter appears in Appendix E.

The same descriptions[1] were used as the basis of peer ratings in this study. The students were asked to read the description and "decide which two girls and two boys in the class best fit this description." Tallies were made of the number of times each student was named.

The instruments for student use were assembled into two booklets. One booklet contained the scale of creative tendency, the social desirability scale, and forms for the ratings of one's peers on these two traits. The second booklet contained the social solidarity scale, the PSM scale, and the corresponding peer rating forms.

The student booklets were administered simultaneously to the two classes during two one-hour periods on successive days. The order in which they were administered was counterbalanced so that one class filled out the first set of questionnaires while the other worked on the second set. The two teachers jointly rated the students on all traits during the first one-hour period.

2) Alternative pattern analysis:[2] Two independent judges rank-ordered the relationships they anticipated among the four traits in this study. Rankings were done on the basis of a priori familiarity with the concepts. The judges had no knowledge of the previously obtained relationships.

3) Regression analyses: Apropos of the multifactorial nature of the PSM scale, Campbell and Fiske (1959) warn that, "Many multitrait-multimethod matrices will show no convergent validation...[if] the trait is not a functional unity [p. 104]." With this in mind, it was decided that some

---

[1] Inasmuch as the same descriptions were employed by peers and teachers, the ratings would be expected to correlate more highly with one another than either one would with scale scores. The derivations reflect the construct to some admittedly imperfect degree.

[2] A paper on pattern analysis by Starr and Seidler is in preparation.

evidence for convergent validity might be gleaned by analyzing the individual subscales of PSM as predictors of various teacher- and peer-rated traits. Regression coefficients were derived from the subscales for both teacher and peer ratings of PSM. A significant multiple correlation between the subscales and the ratings would afford an inference of validity.[1]

RESULTS

Multitrait-Multimethod Analysis: The major purpose of the multitrait-multimethod technique employed here is to establish convergent and discriminant validity for the PSM scale. Campbell and Fiske (1959) have enumerated a number of criteria for the evaluation of multitrait-multimethod matrices. These include:

For convergent validation:

1. The relationships among measures of the same trait assessed by different methods should be high and significant.

For discriminant validation:

2. The relationships among measures of the same trait assessed by different methods should be higher than the relationships between that trait and any other trait assessed by a different method.

3. The relationship among measures of the same trait assessed by different methods should, in general, be higher than the relationships among different traits measured by the same method.

4. The patterning of the trait relationships should be the same irrespective of the method used to assess the traits.

---

[1] It should be noted that this procedure involves making the scales more unitary. Ratings would still be multidimensional (over and above response bias). This would attenuate the likelihood of a high multiple r.

These criteria may now be employed to discuss the validity of the PSM scale.

The multitrait-multimethod matrix for these data is presented in Table 1. As is common practice for such matrices, reliability coefficients

Table 1

are reported on the diagonal. Scale reliabilities are based on data from other studies. No reliability values are available for either teacher ratings or peer ratings, due to time constraints. An estimate of test-retest reliability on these ratings could, of course, be obtained by having teachers and peers make their ratings again.

Evidence for convergent validity requires the demonstration of high and significant relationships among measures of the same trait employing different methods. Correlation coefficients of .30 or higher are needed for significance ($p = .05$) with sample sizes like that in the present study. Of three measures of relationships between pairs of methods of assessing PSM, only one is significant: the relationship between teacher ratings and peer ratings ($r = .591$, $p = .01$). Thus criterion #1 is not met.

The first discriminant validity criterion (#2) requires that the above relationships should be stronger than the relationship of PSM to any other variable which is neither a self-report scale nor a measure of PSM. For each validity coefficient there are six relevant indices. They are the off-diagonal elements of the first row and first column of the heteromethod blocks. With regard to the relationships between teacher ratings and scale scores, three of the six pertinent relationships (see first heteromethod block of figures under "Self-report scales" in Table 1) are stronger than that displayed by the validity coefficient (.16). Two of the six relationships of interest between scale scores and peer

ratings are larger than the relationships between these two measures of PSM (.17). Once again, only the validity coefficient between teacher ratings and peer ratings (.59) demonstrates some degree of validity (according to criterion #2). Here, all six of the studied relationships are weaker than the relationship indexed by the validity coefficient.

It may be of interest to note where the reversals (contraindicators of validity) occur. For the scale-teacher rating validity coefficient, scaled PSM and scaled social solidarity describe a stronger relationship (.45). In addition, teacher-rated PSM has a stronger relationship with scale measures of social solidarity (.34) and social desirability (.17) than with the scale measure of PSM (.16). For the scale-peer rating validator, the scale score relationship cited above (between PSM and social solidarity) once again attenuates the inference of validity. Peer-rated PSM also correlates better with the social solidarity scale (.28) than with the PSM scale (.17).

Another discriminant validity criterion (#3) requires that the validity coefficients should generally be higher than monomethod trait interrelationships. Again the pertinent (monomethod) coefficients are examined separately for each validity coefficient. The teacher rating-scale score validity coefficient (.16) has a lower absolute value than three of the six scale interrelationships and five of the six teacher rating interrelationships. The peer rating-scale score validity coefficient (.17) is lower (in absolute value) than three of the scale coefficients and five of the six peer rating interrelationships. Finally the teacher rating-peer rating validity coefficient (.59) has a lower value than three of the teacher rating monomethod coefficients and one of the six peer rating coefficients. Discriminant validity cannot be established for any of the three validity

coefficients by criterion #3. Once again the teacher rating-peer rating coefficient fares best.

The final criterion (#4) asks that the patterning of trait relationships be the same under all methods. The pattern of relationships is shown in Table 2.

Table 2

The best agreement is between peer ratings and scale scores. This seems, in part, due to the fact that the peers' ratings for different traits are less strongly intercorrelated than teacher ratings. Only the third- and fifth-ranked relationships are reversed. Note, however, that the ranking of the scale score relationships is somewhat arbitrary because of the inclusion of the negative coefficients. Thus, for example, there is less of a relationship between the maturity score and the creative tendency score (the fourth ranked relationship) than between social desirability and creative tendency. Again there is no compelling evidence for validity according to this criterion.

In an attempt to quantify the similarity of patterns obtained by the various methods, the matching formula given by Feller (1968) was applied.[1] The problem was to assign a probability value to the set of patterns obtained. This problem might suggest the use of a coefficient of concordance, but the more general test implied by Feller's (1968) formula, free of erroneous inferences of correlation, appeared to be the better choice. In applying the formula, a double match occurs when one item occupies the same position in two alternative orderings that it occupies in a (third) standard

[1] The application of this formula rests on the assumption that the orderings obtained are not due to sampling error but reflect instead the "true" orderings (i.e., the same rankings that would be obtained in many replications of this investigation).

ordering. The probability of obtaining one or more double matches in two replications of a particular set of six ordered items is .1513 (n.s.). Note that this is a conservative estimate.[1]

Two additional matrices were generated by breaking the original sample into all-boy and all-girl samples. The new matrices did not appreciably change the picture with regard to the validity of the PSM construct, as may be seen in Table 3 where a summary of the multitrait-multimethod matrices is presented. In examining the criterion columns of Table 3, the first three

Table 3

criteria for construct validity are summarized in the column forming the body of the table. Criterion #1 holds that the validity coefficients in the first column must be high and significant. Generally, this is not true for any of the matrices. Criterion #2 requires that all monomethod relationships to the traits of interest (PSM) should be lower than the validity coefficients generated by employing that method. Column 2 contains information on the number of reversals from this desideraturm. Only the teacher rating-peer rating validity coefficient fares well by this criterion. Criterion #3 ideally describes a situation where all monomethod-multitrait relationships (values in the monomethod triangles in Table 1) are weaker than the validity coefficients (Table 3, column 1) generated by that method. Validity can not be demonstrated for PSM by this criterion.

A number of suggestive pieces of information may be gained regarding PSM by examining Table 3 in toto. First, it is clear that the teacher ratings and peer ratings, by virtue of the high validity coefficients which they

[1] Table 2 shows two double matches. Note, however, that the ranking of scale score relationships (the only set containing negative correlation coefficients) is somewhat arbitrary. The ordering is from high positive through low to high negative. Thus, the second match may be an artifact of this particular ranking scheme.

generate, show the best overall validity (i.e., they are most coherent). Secondly, on the basis of two important assumptions, it may be tentatively suggested that validity is more easily obtained with a sample of girls than with a sample of boys. The first assumption involved here is that the validity coefficient obtained between scale scores and peer ratings (r = .399) is non-significant only because of the small sample size. The second assumption may be less well founded. It holds that the differences found between boys and girls are non-trivial (non-chance). The adequacy of this assumption is difficult to assess with the data at hand. Finally, evidence for validity involving scale scores would appear to be more easily obtainable employing a peer (as opposed to teacher) rating method. Again, this finding rests on assumptions similar to the two stated above.

Alternative patterning analysis: The rankings made by two independent judges of the relationships they anticipated among the four traits in this study are shown in Table 4. The probability of the rankings

Table 4

matching one another with one reversal (or less) is small ($p < .011$). Comparison of Tables 2 and 4 reveals that the judges produced orderings which matched the scale score rankings in (at least) two places ($p < .015$).[1] The conceptual orderings also produced double matches for 2 of the 6 teacher-rated relationships ($p < .015$) and 3 of the 6 peer-rated relationships ($p < .002$). Overall, among the five different orderings (three method-generated and two judge-generated), there was (at least) one quintuple match ($p < .005$).[2]

[1] See footnote 1 on P.11

[2] This is based on a generalization of Feller's (1968) formula developed by Alexander J. Seidler

Regression Analyses: Thus far, our analysis has focused on the relationship between total scores on a multi-dimensional trait and ratings of this composite trait as a whole. A source of "noise" in the ratings is the distinct likelihood that different raters may implicitly assign different weights to the components of PSM.

Inasmuch as the focus of this study is on the validation of the PSM construct, the regression coefficients of subscales for both teacher and peer ratings of PSM are shown in Table 5. As noted before, this analysis is predicated on the notion that a significant multiple correlation between the subscales and the ratings affords an inference of validity. The subscales do

Table 5

not yield a significant overall multiple correlation coefficient for either teacher or peer ratings, but two subscales are better-than-chance predictors of teacher-rated PSM. As shown in Table 5, these subscales deal with Self-Acceptance and Social Tolerance. (No other trait rated by teachers or peers was significantly predicted by the subscales). Tangentially, it is worthy of note that the Independence subscale significantly predicts -- and interestingly, negatively predicts -- teacher ratings of social desirability [$t=-2.06, p<.05$].[1]

DISCUSSION

Substantive Results

Before the results are discussed it is important to reiterate the cautionary note sounded at the beginning of this paper. It will be recalled that this paper is in no way viewed as a definitive attempt at validation of the constructs examined here. Rather, the study is conceived as

[1] This finding, while it has a certain intuitive appeal, should not be accorded undue weight. It will be recalled that there were three traits other than PSM and two non-scale methods. Thus, the five subscales were used to generate 30 regression coefficients of which only one was significant. This single regression coefficient could easily represent alpha error.

an exploratory attempt designed to examine potential validation techniques for a new PSM scale. Indeed, the small number of Ss, the truncation of range on a related (age) variable, and the lack of rater training in the study attenuates its proper use for actual validation. With this in mind, findings for the PSM construct are discussed below.

Evidence for the validity of the PSM construct from the multitrait-multimethod matrices generated from these data is generally lacking. This is no doubt partly due to the absence of high and significant validity coefficients for PSM assessed via different methods.[1] Only teacher and peer ratings of PSM yield a significant (non-zero) validity coefficient.

Reasons for this failure (the absence of convergent validation) are not difficult to find. First, the truncation of age-range should have a mitigating effect on the observed correlations. Secondly, the small sample size contributes to a correspondingly unstable correlation coefficient. The net effect is that the coefficients obtained underestimate the true relationships; and that all of the coefficients (low or high) are relatively poor estimates of the actual correlation. Finally, Campbell and Fiske (1959) suggest that low validity coefficients emanate from two additional and distinct situations:

(1) one or more of the methods is not measuring the trait

(2) the trait is non-unitary (i.e., not factorially "pure")

Inasmuch as PSM is a multidimensional construct the likelihood of finding high convergent relationships is small. Moreover, it appears that the methods did clearly differ in their relationship with one another. This is not particularly surprising in light of the relative degree of similarity among methods. It will be recalled that the stimuli forming the basis for ratings were virtually identical for teachers and peers. These stimuli, by

[1] It will be recalled that two of the other validity criteria require that other coefficients be smaller in magnitude than the validity coefficients.

the nature of their construction, were more similar to one another than either stimulus set was to the scale-stimulus. Thus, notably stronger relationships appear in the heterotrait-heteromethod block involving the two rating methods than in either of the other heterotrait-heteromethod blocks. The validity diagonal for the rating (heterotrait-heteromethod) block reflects the same general pattern of relatively stronger relationships. The methods also differed in their ability to distinguish traits. Monomethod relationships are notably higher for both of the rating methods than for the scales. This implies that teachers and peers tend to "level" rather than sharpen distinctions among the traits of individuals when compared with scales. The suggested interpretation has some intuitive appeal. It may be argued that people tend to make finer distinctions about themselves (i.e., via self-report scales) than about others. Certainly "leveling" appears to be an important mode of human information processing. This may be especially true in situations which minimize ego-involvement. Alternatively, scales would probably provide better means for making distinctions (in view of their more complex structure) than would ratings. At any rate, the absence of high convergence in a multitrait-multimethod matrix undercuts any attempt at establishing validity via that matrix.

Finally, while evidence for validity is quite scant in this study, some results bear further examination. The predicted patterning of results by two independent raters yielded patterns similar to the one which was obtained for scale score relationships beyond a chance level. Indeed, there was a tendency toward significant similarity among the patternings obtained for the three methods. Also noteworthy was the fact that two of the five subscales yielded significant regression coefficients in predicting

teacher ratings of PSM.[1] The remaining rated traits were also examined via a regression analysis. While PSM ratings were not significantly predicted by the subscales taken together, two subscales individually predicted teacher-rated PSM.

One further caution is in order with regard to the results of the regression analysis. Despite the fact that the subscales constitute the PSM self-report measure, thus implying good a priori grounds for their use as predictors of rated PSM, the statistical fact remains that, overall, three of 40 (5 subscales x 2 rating methods x 4 traits) regression coefficients were significant. Once more the evidence should only be regarded as suggestive. Alpha error remains a plausible alternative explanation of these findings. A study that was not exploratory would, of course, examine only the regression coefficients which had an a priori relationship to a particular trait.

RECOMMENDATIONS FOR FUTURE ATTEMPTS AT VALIDATION

The exploratory nature of this research has been referred to throughout the paper. As has been noted, the PSM scale is currently undergoing extensive revision, and the revised scale will become the basis for the major attempts at validation. As had been anticipated, the present investigation has implications for the procedures in subsequent studies.

First, it should be noted that the multitrait-multimethod studies have gained currency as the sine qua non of validation efforts. Yet, it is not widely realized that some authors have found difficulties with the method. Campbell (1960) has outlined and argued against a number of the problems which have been raised. A recent paper (Wallace, 1965) came to grips with the issue of criteria. There Wallace raises a cogent question:

---

[1] It will be recalled, however, that the overall ability of the subscales to predict rated PSM was nonsignificant.

"Would a systematic study of the intercorrelations of...tests...valid for rating criteria reveal....only....a 'good guy' or 'impressive-fellow' factor"[p.412, italics added]. In the present study, higher correlations between teacher and peer ratings than between ratings and scale scores might suggest just such a factor (in view of the minimal instructions to the raters).

A provocative paper by Ebel (1961) subjects the concept of validity to careful study. He observes that psychology, in its emulation of "harder" sciences, has failed to note the lack of concern in these other disciplines with validity. This is so despite the fact that some quantified properties lack consistency when measured by different procedures. Moreover, Ebel submits that newer methods in hard sciences are not justified on the basis of validity, but rather on the basis of superior reliability. He also suggests the ludicrous nature of attempting to check scores on new (and hopefully better) instruments against less "good" measures of the trait of interest. In dealing with complex traits or criteria, it is clear that attention must be paid to interrelationships among measures. Still, Ebel cogently raises the question of whether we are being overzealous in our pursuit of validity. Ebel's remarks address criterion and prediction problems. In the important preliminary procedures for construct validation, the domain of the construct must be clearly articulated (Nunnally, 1967). These procedures essentially involve examination of relationships among measures which are potentially expected to be within the domain. Given that some measures fail inclusion in the domain, such measures will be less "good" indicants of the construct. If such measures are included in a multi-trait-multimethod study, one might be in the position of attempting to establish construct validity using a relatively new measure that relates to

the construct better than the other measures included in the study.

Given the complexity of traits and methods in this study, it should not be at all surprising to find monomethod coefficients which describe stronger relationships than monotrait coefficients. Relatively high method variance alone can not be taken as an indictment of the adequacy of the method. One can show evidence of some construct validity despite high method variance. This should not be taken to mean that multitrait-multimethod studies are valueless. It does suggest that we may be according such studies undue weight. The critical question prior to validation must be what end will the study serve. If the end is prediction, the criterion must be of prime concern. If the end is understanding, intercorrelations may be more informative. However, successful multitrait-multimethod studies may require considerable investment to produce this information. Reduction of method variance for complex traits may require a substantial effort to align the methods. The upshot of such effort may be little more than the production of a number of interchangeable methods -- some never to be employed again. If this is a desideratum, the findings of the present study are suggestive of mechanisms improving the quality of multitrait-multimethod matrices in future studies like this one.

We have discussed earlier the difficulties involved in validating measures of multidimensional constructs. In future studies of PSM, individual subscales should probably become the focus of validation efforts, and careful descriptions of the trait assessed by the subscale should be created for use by the rater. Raters need to be cautioned against "halo effects" and their operation in human judgment. And, finally, the evidence that girls' PSM scores show somewhat better validity must be thought through and procedures devised to determine why a sex difference occurs.

It is clear that construct validity cannot be purchased cheaply. Programmatic research in establishing the domain of the constructs should precede multitrait-multimethod studies. The articulation of the domain of a construct provides useful information on alternative methods which might be employed. Absence of such information leaves open a number of alternative explanations of obtained results. Thus, in the present study, any one or more of the measures employed may not be part of the domain of the construct. It could well be argued that multitrait-multimethod studies are useful as the capstone of programs of research on constructs; and that their early application may be of limited utility.

Clearly, with PSM there may be more important goals to examine. Does a higher score indicate effective functioning? In what ways is this true? For example, is a high Independence score associated with better individual functioning? Once such hypotheses are developed, it is critical to evolve behavioral criteria where possible, in order to obviate exclusive reliance on ratings of behavior by others.

REFERENCES

Beers, J., Russell, N., and Burson, W. Manual for interpreting school reports. Harrisburg, Pa.: Pennsylvania Department of Education, 1971.

Campbell, D. T. Recommendations for APA test standards regarding construct, trait, or discriminant validity. American Psychologist, 1960, 15, 546-553.

Campbell, D. T., and Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.

Crowne, Douglas P., and Marlowe, D. The Approval Motive: Studies in Evaluative Dependence. New York: Wiley, 1967.

Ebel, R. L. Must all tests be valid? American Psychologist, 1961, 16, 640-647.

Feller, W. An introduction to probability theory and its applications. (3rd Ed.) Vol. 1, New York: Wiley, 1968.

Greenberger, E. Psychosocial maturity or social desirability? Report No. 131, Baltimore, Maryland. Center for Social Organization of Schools, The Johns Hopkins University, 1972.

Greenberger, E. and Marini, M.M. Black-white differences in psychosocial maturity: a further analysis. Report No. 136. Baltimore, Maryland. Center for Social Organization of Schools, The Johns Hopkins University, 1972.

Greenberger, E. and Sørensen, Aa. B. Toward a concept of psychosocial maturity. Report No. 108. Baltimore, Md. Center for Social Organization of Schools, The Johns Hopkins University, 1971.

Greenberger, E., Campbell, P., Sørensen, Aa. B., and O'Connor, J. Toward the measurement of psychosocial maturity. Report No. 110. Baltimore, Md. Center for Social Organization of Schools, The Johns Hopkins University, 1971.

Nunnally, J. C. Psychometric theory. New York: McGraw-Hill, 1967.

Starr, B. J., Greenberger, E., Seidler, A. J., Marini, M. M., Sørensen, Aa. B., Campbell, P. and O'Connor, J. Black-white differences in psychosocial maturity. Report No. 127. Baltimore, Md. Center for Social Organization of Schools, The Johns Hopkins University, 1972.

Wallace, S. R. Criteria for What? American Psychologist, 1965, 20, 411-417.

Please read the following description of a person carefully. Then decide where all of your students belong relative to one another and place them under the appropriate numbered category along the "trait line."

DESCRIPTION: This person
- --has ideas which are often very unusual or surprising
- --has talents (for example, writing or telling stories, painting, or carving) that many other children don't have
- --sometimes surprises the class with good, but unexpected answers
- --has his or her own way of doing things

TRAIT LINE

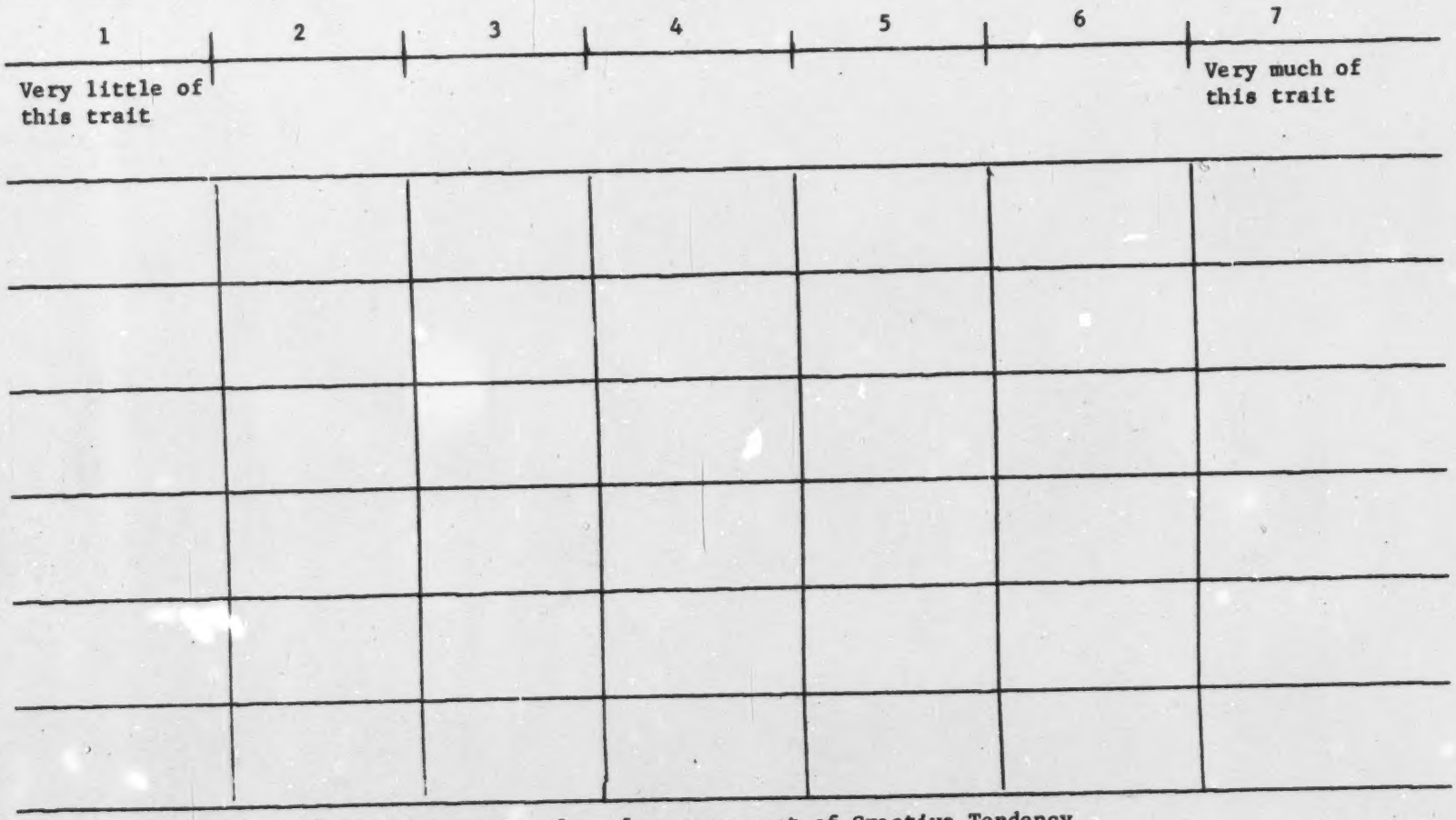

Fig. 1. Teacher-rating form for assessment of Creative Tendency

Table 1

Multitrait-Multimethod Matrix for PSM, Social Solidarity, Social Desirability, and Creative Tendency

		Self-report scales				Teacher ratings				Peer ratings			
		PSM	Soc. Sol.	Soc. Des.	Cr.	PSM	Soc. Sol.	Soc. Des.	Cr.	PSM	Soc.	Soc.	Cr.
Self-report scales	PSM	.81[a]											
	Soc. Sol.	.45	.78[b]										
	Soc. Des.	.02	.30	.88[c]									
	Creat.	.01	-.04	-.21	.82[d]								
Teacher ratings	PSM	.16	.34	.17	.08	—							
	Soc. Sol.	.01	.25	.22	.13	.63	—						
	Soc. Des.	-.27	.11	.08	.15	.34	.68	—					
	Creat.	.13	.19	.03	.07	.73	.52	.15	—				
Peer ratings	PSM	.17	.28	.09	.23	.59	.47	.26	.57	—			
	Soc. Sol	-.06	.18	.29	.08	.47	.56	.37	.45	.76	—		
	Soc. Des.	-.15	.11	.19	.27	.28	.42	.31	.16	.37	.47	—	
	Creat.	-.07	.10	.20	.01	.44	.42	.17	.53	.45	.46	.13	—

Note -- Coefficients on the diagonal are reliability values. No such values were obtained for teacher or peer ratings.

a. Source for this reliability value is Greenberger et al (1971)
b. This reliability value was obtained from an as yet unpublished study by Greenberger
c. This reliability value was reported for the unmodified Social Desirability scale by Crowne & Marlowe
d. Source for this reliability value is Beers, Russell, and Bunson (1971)

Table 2

Patterning of Monomethod Relationships: Criterion 4

	Methods		
Rank	Scale Scores	Teacher Ratings	Peer Ratings
1	PSM-SS(.45)	PSM-CT(.72)	PSM-SS(.76)
2	SS-SD(.30)	SS-SD(.68)	SS-SD(.47)
3	PSM-SD(.02)	PSM-SS(.63)	SS-CT(.46)
4	PSM-CT(.01)	SS-CT(.52)	PSM-CT(.45)
5	SS-CT(-.04)	PSM-SD(.36)	PSM-SD(.37)
6	SD-CT(-.21)	SD-CT(.15)	SD-CT(.13)

Note: - PSM = psychosocial maturity; SS = social solidarity; SD = social desirability; CT = creative tendency.

Table 3

Summary of Three Multitrait-Multimethod Matrices Examining Validity Evidence for the PSM Construct

Methods	I. Validity Coefficients	Criteria No. of reversals (invalidating coefficients) of six possible invalidators II.[a]	III.[b]
Main sample (Boys and Girls Combined), n = 47			
1. Scale - teacher rating	.159	3	3 5
2. Scale - peer rating	.172	2	3 5
3. Teacher rating - Peer rating	.591*	0	3 1
Girls, n = 21			
1. Scale - teacher rating	.169	5	4 4
2. Scale - peer rating	.399	1	2 3
3. Teacher rating - Peer rating	.736*	0	0 3
Boys, n = 26			
1. Scale - teacher rating	.111	3	3 6
2. Scale - peer rating	-.085	5	4 6
3. Teacher rating - Peer rating	.462*	1	4 2

a. No. of row and column coefficients demonstrating stronger relationships than the validity coefficient.

b. No. of monomethod coefficients demonstrating stronger relationships than the validity coefficient.

* $p < .01$

Table 4

Predicted Patterning of Relationships

Rank	Rater #1	Rater #2
1	PSM-SS	PSM-SS
2	SS-SD	SS-SD
3	PSM-CT	PSM-SD
4	PSM-SD	PSM-CT
5	SS-CT	SS-CT
6	SD-CT	SD-CT

Note - PSM = Psychosocial Maturity; SS = Social Solidarity; SD = Social Desirability; CT = Creative Tendency

Table 5

Raw Regression Coefficients for Regression of Teacher Ratings and Peer Ratings of PSM on the PSM subscales

PSM Subscale	Ratings of PSM: Teacher Ratings Regression Coefficient	Teacher Ratings Standard Error	Peer Ratings Regression Coefficient	Peer Ratings Standard Error
1. Self-Acceptance	0.45**	0.18	0.49	0.69
2. Openness to Change	-0.04	0.11	-0.15	0.41
3. Independence	-0.10	0.12	0.82	0.44
4. Identity	0.01	0.24	-1.26	0.93
5. Social Tolerance	0.52*	0.22	0.79	0.85

* $p < .05$

** $p < .02$

APPENDIX A

Psychosocial Maturity

The questionnaire format for the 54 items of the psychosocial maturity scale is given below. The asterisk indicates which response or responses are "mature."

DIRECTIONS: Please mark each statement in the following way: If the statement describes how you usually feel, check "Like me". If the statement does not describe how you usually feel, check "Unlike me". There are no right or wrong answers.

Subscale: Self-Esteem	Like me	Unlike me
1. Luck decides most things that happen to me.		*
2. Someone always has to tell me what to do.		*
3. It takes me a long time to get used to anything new.		*
4. I'm popular with kids my own age.	*	
5. If I work hard, I can be what I want to be.	*	
6. If I stick to something long enough, I can make it work.	*	
7. If I work hard, I can get a good job.	*	
8. If I have something to say, I usually say it.	*	
9. There isn't much of a chance for a person like me to succeed in life.		*
10. If I work at something long enough, I will succeed.	*	

SOME PEOPLE SAY THAT IN ORDER TO KEEP UP IN A FAST MOVING WORLD IT WILL BE IMPORTANT TO KEEP LEARNING AND STUDYING ALL DURING OUR LIVES. OTHER PEOPLE SAY THAT ONCE A PERSON FINISHES SCHOOL, HE SHOULD BE ABLE TO HANDLE ANYTHING THAT COMES ALONG.

DIRECTIONS: Check the one column that best describes you. If you wish to change an answer, erase completely your first mark.

DO YOU THINK YOU WILL HAVE TO KEEP LEARNING AND STUDYING (IN OR OUT OF SCHOOL) IN ORDER TO:

Subscale: Openness to Change	It will be very important	It will be quite important	It will be somewhat important	It will not be important	I can not say
11. Make good decisions in voting.	*	*			

DIRECTIONS: TRY TO PLACE YOURSELF IN THE FOLLOWING SITUATION:

BEFORE SCHOOL BEGAN ONE MORNING, YOU AND A GROUP OF CLASSMATES WERE HAVING A TALK ABOUT THE YEAR 1989 AND YOU WERE TRYING TO IMAGINE YOURSELVES GOING TO SCHOOL IN 1989. ITEMS 12 to22 ARE SOME OF YOUR CLASSMATES IDEAS.

	I accept it	I can not say	I do not accept it
12. There will be no marks and no report cards. Pupils will talk over their work with their teachers as often as they like.	*		
13. Schools will be open 24 hours each day. Pupils can use the building at any time.	*		
14. Pupils will work with teachers alone or in small groups.	*		
15. All the latest and best reading materials will be quickly available through a computer.	*		
16. There will be TV, movies, records, and tapes which pupils can use by themselves.	*		

	I accept it	I can not say	I do not accept
17. Pupils will have the use of a computer for arithmetic and many other things.	*		
18. There will be quiet places to learn and study on one's own.	*		
19. To learn about the people and the language of another country, pupils will spend some time living in foreign countries.	*		
20. To learn about different people in this country, pupils will spend time living in different sections of the United States.	*		
21. Pupils will learn in factories, laboratories, hospitals, museums, theaters and offices. Pupils will visit these places if they wish to learn about them and about the people in them.	*		
22. Pupils will talk with others all over the world by way of satellite.	*		

DIRECTIONS: READ EACH STATEMENT, THINK ABOUT WHAT YOU HAVE DONE IN THE PAST IF YOU WERE EVER FACED WITH THE SITUATION AND MARK YOU ANSWER BY CHECKING THE COLUMN THAT FITS BEST. IF YOU NEVER FACED THE SITUATION, TRY TO IMAGINE WHAT YOU WOULD DO. DON'T SPEND TOO MUCH TIME ON ANY ONE ITEM.

Subscale: Independence	Always	Most of the time	Some-times	Very seldom	Never
23. I like to earn my own money when I can.	*	*			
24. A man should work and earn his own living if he can.	*	*			
25. A man should vote the same way his friends do.				*	*
26. People should not be allowed to say what they think.				*	*

DIRECTIONS: THE ITEMS BELOW ARE STATEMENTS ABOUT OCCUPATIONAL CHOICE AND WORK. OCCUPATIONAL CHOICE MEANS THE KIND OF JOB OR WORK THAT YOU THINK YOU WILL PROBABLY BE DOING WHEN YOU FINISH ALL OF YOUR SCHOOLING. IF YOU AGREE OR MOSTLY AGREE WITH THE STATEMENT, PUT A CHECK IN THE SPACE HEADED TRUE. IF YOU DISAGREE OR MOSTLY DISAGREE WITH THE STATEMENT, PUT A CHECK IN THE SPACE HEADED FALSE.

Subscale: Independence		True	False
	27. I plan to follow the line of work my parents suggest.	____	*
	28. I'm not going to worry about choosing an occupation until I'm out of school.	____	*
	29. Your parents probably know better than anyone else which occupation you should enter.	____	*
	30. Why try to decide on a job when the future is so uncertain.	____	*
	31. I seldom think about the job I want to enter.	____	*
	32. It doesn't matter which job you choose as long as it pays well.	____	*
	33. You can't go very far wrong by following your parents' advice about which job to choose.	____	*
	34. Whether you're interested in a particular kind of work is not as important as whether you can do it.	____	*
	35. You get into an occupation mostly by chance.	____	*
	36. It's who you know, not what you know, that is important in a job.	____	*
	37. When it comes to choosing a job, I'll make up my own mind.	*	____
	38. I have little idea of what working will be like.	____	*
	39. Choose an occupation, then plan to enter it.	*	____
	40. I really can't find any work that appeals to me.	____	*
	41. Choose a job in which you can be famous.	____	*
	42. The most important part of work is the pleasure which comes from doing it.	*	____
	43. Why worry about choosing a job when you don't have anything to say about it.	____	*

Subscale: Identity

	True	False
44. I don't know how to go about getting into the kind of work I want to do.	____	* ____
45. I don't know what courses I should take in school.	____	* ____
46. I know very little about the requirements of jobs.	____	* ____
47. I can't understand how some people can be so set about what they want to do.	____	* ____

DIRECTIONS: READ EACH QUESTION CAREFULLY AND DECIDE HOW YOU FEEL ABOUT IT. THERE ARE FIVE POSSIBLE ANSWERS TO CHOOSE FROM. BE SURE TO ANSWER EACH QUESTION. CHECK ONLY ONE SPACE FOR EACH QUESTION.

Subscale: Social Tolerance

	I would like it	I would not mind it	I would rather not	I would dislike it	I can not sa
48. How would you feel about sitting in class next to a person whose skin color is different from your own?	* ____	* ____	____	____	____
49. How would you feel about having as a best friend a person whose ideas about God are very different from your own?	* ____	* ____	____	____	____
50. How would you feel about playing on the same team with a person whose ideas about God are very different from your own?	* ____	* ____	____	____	____
51. How would you feel about sitting in class next to a person whose ideas about God are very different from your own?	* ____	* ____	____	____	____
52. How would you feel about having as best friend a person whose family is much poorer than yours?	* ____	* ____	____	____	____
53. How would you feel about playing on the same team with a person whose family is much poorer than yours?	* ____	* ____	____	____	____
54. How would you feel about sitting in class next to a person whose family is much poorer than yours?	* ____	* ____	____	____	____

APPENDIX B

Social Solidarity

Please read each item carefully. Place a check in the space which best describes how you feel. Do not leave any items unanswered.

	Mostly agree	Mostly disagree
1. I like the people in my neighborhood.	*	
2. I tend to forget about the feelings of others when I'm doing something that is important to me.		*
3. A person is responsible only for the well-being of his family, relatives, and close friends.		*
4. Students should take part in the group activities of their class.	*	
5. I do not like other kids to ask me to help them with their homework.		*
6. Most people pretend their troubles are worse than they are to get sympathy.		*
7. It is more satisfying to work for a purely personal goal than to work for a goal held by a group you belong to.		*
8. Most people like to help others.	*	
9. A person should not be respected for something he did, if it interfered with the well-being of others.	*	
10. Sometimes it's necessary to push people around a little to get what you want.		*
11. I do not like to teach other kids how to do things.		*
12. I like most people I meet.	*	
13. It's natural for each person to think that his family deserves to have things better than any other.		*
14. If you really enjoy doing something, you should do it even if it causes problems for other people.		*
15. It's none of my business if my neighbors are in trouble and need help.		*
16. I don't see much point in trying to do things for others unless they can do you some good later on.		*

		Mostly agree	Mostly disagree
17.	I would enjoy working with others to plan a class outing.	*	
18.	I would not want to pay taxes to run schools if I did not have children.		*
19.	I'm not the sort of person who goes out of my way to help others.		*
20.	I would not mind serving on a committee for my school.	*	
21.	People who ask for help are giving up their pride.		*
22.	I admire a good follower as much as I admire a good leader.	*	
23.	I'd like to be part of a family where the parents do not expect their children to do jobs around the house.		*
24.	Once you start doing favors for people, they'll just walk all over you.		*
25.	When you get right down to it, no one is going to care much what happends to you.		*
26.	I don't like to lend money to my friends.		*
27.	When the chips are down, I don't have any friends I can really count on.		*
28.	I don't see why my parents should pay taxes to provide food for the poor people of other states.		*
29.	When a new person moves into my neighborhood, I try to be friendly.	*	
30.	If I had to choose between helping raise money for a neighborhood project and enjoying my own free time, I'd probably keep my freedom.		*
31.	I don't like to lend my extra pencils or pens in school.		*
32.	If I had the choice of <u>working with</u> somebody so each of us could get part of a prize, or <u>competing</u> against him so one would get it all, I'd compete.		*

APPENDIX C

NAME _____

GRADE _____ TEACHER _____

Social Desirability

		True	False
1.	I always do the right thing.	*	
2.	I never hesitate to go out of my way to help someone in trouble.	*	
3.	It is sometimes hard for me to go on with my work if I am not encouraged.		*
4.	I have never disliked anyone.	*	
5.	I have sometimes had doubts about whether I can succeed in life.		*
6.	I sometimes feel angry when I don't get my way.		*
7.	I am always careful about how I dress.	*	
8.	My table manners at home are as good as when I eat out in a restaurant.	*	
9.	If I could get into a movie without paying and be sure I was not seen, I would probably do it.		*
10.	A few times I have given up doing something because I thought I didn't have the ability to do it.		*
11.	I like to gossip at times.		*
12.	There have been times when I did not feel like doing what my teachers wanted me to do, even though I knew they were right.		*
13.	No matter who I'm talking to, I'm always a good listener.	*	
14.	I can remember "playing sick" to get out of something.		*
15.	There have been times when I took advantage of someone.		*

		True	False
16.	I'm always willing to admit it when I have made a mistake.	*	
17.	I always try to do what I tell other people is the right thing to do.	*	
18.	I don't find it very hard to get along with rude, disagreeable people.	*	
19.	I sometimes try to "get even" with someone I'm angry at, rather than "forgive and forget."		*
20.	When I don't know something, I don't mind saying so at all.	*	
21.	I am always polite, even to people who aren't very nice.	*	
22.	At times I have really insisted on having things my own way.		*
23.	There have been times when I felt like smashing things.		*
24.	I would never think of letting someone else be punished for something I did.	*	
25.	I am never annoyed at being asked to do a favor for someone who has done a favor for me.	*	
26.	I have never been annoyed when people expressed ideas that are very different from my own.	*	
27.	There have been times when I was jealous of other people's good luck.		*
28.	I have almost never felt like telling someone that I don't think much of him.	*	
29.	I am sometimes irritated by people who ask me to do them favors.		*
30.	I have never felt that I was punished without a reason.	*	

		True	False
31.	I sometimes think when people have something bad happen to them, they only got what they deserve.	____	*____
32.	I have never purposely said something that hurt someone's feelings.	*____	____
33.	I tell a little lie sometimes.	____	*____

APPENDIX D

Creative Tendency

SECTION F

DIRECTIONS: READ EACH STATEMENT CAREFULLY. BLACKEN THE CIRCLE WHICH BEST TELLS HOW YOU FEEL ABOUT THE STATEMENT. FOR EACH ITEM BLACKEN ONLY ONE CIRCLE.

Strongly disagree — Disagree — Uncertain — Agree — Strongly agree — Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ

1. If the last page of a book is missing, the book is not worth reading ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
2. I would like to make up a new song ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
3. New words are fun to learn ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
4. T.V. news shows are boring ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
5. I would like to learn a new game even if I lose at it ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
6. The best lessons contain only one idea ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
7. New places are fun to visit ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
8. I like to really do more than the teacher asks for ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
9. I think that it is foolish to borrow other people's ideas ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
10. The best toys are the kind that can be taken apart ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
11. I like to meet new classmates ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
12. Creating new words is dumb ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
13. It might be fun to try new games ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
14. Learning how to do things is more important than getting excellent marks ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
15. I like to try new things ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
16. I like to make things without following directions ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
17. Pictures of grass should be colored only green ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ

Strongly disagree — Disagree — Uncertain — Agree — Strongly agree — Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ

18. Only adults can run a house ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
19. I think I could make up stories as good as those in books ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
20. You have to be grown up to think up a really good idea ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
21. I don't like changes ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
22. It is important to get the right answer even if I don't know how I got it ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
23. I would rather play a game I know than learn a new one ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
24. I enjoy learning how to do something in a new and different way ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
25. The best way is the old way ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
26. It would be fun to take a picture through a goldfish bowl ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
27. I would rather copy poems from a book than make them up ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
28. It is not much fun to visit new places ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
29. I want to find out things that nobody else knows ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
30. Games are not fun if you lose ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
31. The more pieces in a puzzle, the better I like it ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
32. The best friends are the ones who like the same things as I do ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
33. Lessons that have a lot of different ideas in them are really good ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
34. I don't like to learn new words ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
35. News shows on T.V. are interesting to watch ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
36. A good drawing has to look like the real thing ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
37. I like to speak to the class and answer questions about my talk ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
38. Singing a song that nobody else knows is silly ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
39. If the last page of a story is missing, you should make up the ending ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
40. Pictures of grass could be painted any color ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
41. Making up stories is silly ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
42. It would be fun to draw a picture while standing on your head ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
43. New kids are not fun to meet ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
44. It would be a waste of time to take a photograph through a fishbowl ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ
45. I think that it is foolish to copy other people's work ..... Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ

APPENDIX E

THE JOHNS HOPKINS UNIVERSITY · BALTIMORE, MARYLAND 21218

CENTER FOR SOCIAL ORGANIZATION OF SCHOOLS
301-366-3382

TO THE TEACHER:

The questionnaire being given to your class is part of a survey being done by people in the Pennsylvania State Department of Education and the Johns Hopkins University. This survey is part of an ambitious study of some important non-academic traits in human development. Perhaps the clearest picture of where a person stands on a trait emerges from looking at how an individual is seen by others as well as how he sees himself. Children's teachers and peers are both valuable "observers." On the next pages, we are asking you to help us develop this method of looking at aspects of a child's personality. Obviously, the success of this endeavor depends a great deal on your candor and your willingness to make clear distinctions among your students. Your judgements will be treated with absolute confidentiality.

The next 3 sheets describe different traits. The trait descriptions are exactly the same as those given to your students. We have done this purposely in order to maximize the overlap in content of the traits that both you and your students will be judging. Please rate all of your students on the trait. Rate the students relative to one another according to where they belong with regard to the numbered categories along the "trait line." If you rate them relative to one another you should be able to use all of the numbered categories. Please write the names of the students at a particular point on the trait line directly under the number. If you work in pencil you will be able to readjust your discriminations as you think of more students. We are grateful for your conscientiousness and cooperation. Thank you.

B. James Starr, Ph.D.
Ellen Greenberger, Ph.D.